

Part of 09/990,437 page 9

Query Match 16.8%; Score 882; DB 2; Length 2186;
Best Local Similarity 27.9%; Pred. No. 1.2e-70;
Matches 295; Conservative 147; Mismatches 338; Indels 276; Gaps 43;

Qy	5	EMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSR-RAFQELVLEPAQRRARLEGLR	63
Db	1231	EQKKYESEESVSKGSW----QKTVNNNQQSLFQRLDFKSKDISKIAADITQAVSLSQGIE	1286
Qy	64	YTAVLKQQATQHSM---ALLHWGALWRQLASPCGAWALRDTPI---PRWKLSSAETYSRM	117
Db	1287	RKKVIQHIRGMYKVDLSASRHWQECIQQLTHDRAVWY---DPIYYPTSWQLDPTEGPNRE	1343
Qy	118	RLKL-----VPNHHFDPHLEASALRD-NLGEVPLTPTEEASLPLAVTKEAKVSTPPELL	170
Db	1344	RRRLQRCYLTI PNKYL-----LRDRQKSEGVLRP-----PLSYLFEDKTHSSFSST	1389
Qy	171	QEDQLGEDELAELETPMEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFY-D	229
Db	1390	VKDKAASESIRVNRRCISVAPSRETAGELLL-GKC-----GMYFVED	1430
Qy	230	GSTERVETEEGIG----YDFRRPLAQLREVHLRRENLRRSALELFFIDQANYFLNFPCKV	285
Db	1431	NASDAVESSSLQGELEPASFSWTYEEIKEVHRRWWQLRDNAVEIFLTNGRTL LLLAF----	1486
Qy	286	GTPPVSSPSQTPRPQPGPIPPHTQVRNQVYSWLLRLRPSPQGYLSSRSPQ--EMLRASGL	343
Db	1487	-----DNNKVRDDVY-----QSILTNNLPNLLEYGNITAL	1516
Qy	344	TQKWVQREISNFEYLMQLNTIAGRTYNDLSQYPVFPWVLQDYVSPTLDLSNPAVFRDLSK	403
Db	1517	TNLWYSQGQITNFEYLTNLKHAGRSFNDLMQYPVFPFILSDYVSETLDLNDPSIYRNLSK	1576
Qy	404	PIGVVNPKHAQLVREKYESFE-----DPAGTIDKFHYGTHYSNAAGVMHYLIRVE	453
Db	1577	PIAVQYKEKEDRYVDTYKYLEEEYRKGAREDDPMPPVOPYHYGSHYSNSGTVLHFLVRMP	1636

Qy 454 PFTSLHVQLQSGRFDCSDRQFHSAVAAWQ-ARLESPADVKELIPEFFYFPDFLENQNGFD 512
 ||| : : | || || || | : : || ||||| ||| : || : |||
 Db 1637 PFTKMFLAYQDQSFDPDRTFHSTNTTWRLSSFESMTDVKELIPEFFYLPEFLVNREGFD 1696

Qy 513 LGCLQLTNEKVGDVVLPWA-SSPEDFIQQHRQALESEYVSAHLHEWIDLIFGYKQRGPA 571
 | | : | | ||| : | || ||||| : || : : ||| : ||| : | |
 Db 1697 FGVQRQ-NGERVNHVNLPPWARNDPRLFILHRQALESDHVSQNICHWIDLVFYKQKGKA 1755

Qy 572 AEEALNVFYCYTYEGAVDLHDVTDERERKALEGIISNFGQTPCQLLKEPHPTRLSAE--- 628
 : : ||| : || : : | | : ||| : | : ||| || | : | :
 Db 1756 SVQAINVFHPATYFG-MDVSAVEDPVQRRALETMIKTYGQTPRQLFHTAHASRPGAKLNI 1814

Qy 629 -----EAAHRLAR---LDTNSPSIFQHLDELKAFFAEVTVSASG----- 664
 : | | | : || : || : | | |
 Db 1815 EGELPAAVGLLVQFAFRETREPVEVTHPSPLSWIKGLK--WGEYVGSPSAPVPVVCFSQ 1872

Qy 665 -----LLGTH-----SWLPYDRNIS 679
 || | : || | ||
 Db 1873 PHGERFGSLQALPTRAICGLSRNFCLLMTYNKEQGVRSMNNTNIQWSAILSW-GYADNIL 1931

Qy 680 NYFSFSKDPTMG---SHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGH-----WD 730
 | : : | : : | || || ||| : :
 Db 1932 RLKSKQSEPPINFIQSSQHQVTSCAWV-----PDSCQLFTGSKCGVITAYT 1978

Qy 731 GSLRVTALPRGKLLSQLSC--HLDVVTCLALDTCGIY--LISGSRDTCMVW---RLLHQ 783
 | : : ||| | : : | | : || ||| ||| : || :
 Db 1979 NRLTSSTPSEIEMESQMHLGYHTEETGLCV--CKPYSVMISVSRDGTICVWDLNRLCY- 2035

Qy 784 GGLSVGLAPKPVQVLYGHGAAVSCVA-----ISTELDMAVSGSE-----DGTVIIHT 830
 || | || : | : : ||| | | || : | : : |
 Db 2036 -----VQSLAGHKSPVTAVSASETSGDIATVCDSAGGGSDLRLWTVNGDLVGH- 2083

Qy 831 VRRGQFVAAL---RPLGATFPGPFIHFLALGSEGQIVVQSSAWE-----RPG 873
 | : : : : | | : | : | | || | : : |
 Db 2084 VHCREIICSVAFSNQPEGVS----INVIAGGLENGIVRLWSTWDLKPVREITFPKSNKPI 2139

Qy 874 AQVTYSL---HLYSVNG-----KLRASLPL 895
 : || | ||| : | : |||
 Db 2140 ISLTFSCDGHLYTANSEGTVIAWCRKDQQRVKLPM 2175

RESULT 3

US-09-396-540-2

; Sequence 2, Application US/09396540

; Patent No. 6310182

; GENERAL INFORMATION:

; APPLICANT: Kaplan, Jerry

; APPLICANT: Perou, Charles

; APPLICANT: Moore, Karen

; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS

; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

```

; Sequence 1, Application US/08822445
; Patent No. 5952223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6830 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...6558
; OTHER INFORMATION:
US-08-822-445-1

```

```

Query Match          5.0%; Score 176.8; DB 2; Length 6830;
Best Local Similarity 52.9%; Pred. No. 5.5e-31;
Matches 475; Conservative 0; Mismatches 387; Indels 36; Gaps 3;

```

```

Qy      1312 CTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACCTTCGAGTACTTGATGCA 1371
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4539 CACCGCTCTGACAAACCTGTGGTATTCTGGACAAATTACCAATTTTGAATATTTGACTCA 4598

Qy      1372 ACTCAACACCATTCGCGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTCCCTG 1431
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4599 TTAAACAAGCATGCGGGCCGGTCCTTCAATGATCTCATGCAGTACCCGGTGTTCCTT 4658

Qy      1432 GGTCTGCAGGACTACGTGTCCCAACCCTGGACCTCAGCAACCCAGCCGTCTTCCGGGA 1491
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4659 CATCCTTTCTGACTATGTTAGTGAGACTCTTGACCTCAATGATCCATCTATCTACAGAAA 4718

Qy      1492 CCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCAGCTCGTGAGGGAGAAGTA 1551
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4719 CCTATCTAAGCCTATAGCTGTGCAGTATAAAGAAAAAGAACCGTTACGTTGACACATA 4778

Qy      1552 TGAAAGCTTTGAG-----GACCCAGCAGGGACCAT 1581
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4779 CAAGTACTTGAGAGGAGGAGTATCGCAAGGGAGCTCGAGAGGATGACCCCATGCCTCCTGT 4838

Qy      1582 TGACAAGTTCCACTATGGCACCCACTACTCCAATGCAGCAGGCGTGATGCACTACCTCAT 1641
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4839 GCAACCCTACCACTATGGCTCCCACTACTCCAACAGCGGCACCGTGCTCCACTTCCTGGT 4898

Qy      1642 CCGCGTGGAGCCCTTACCTCCCTGCACGTCCAGCTGCAAAGTGGCCGCTTTGACTGCTC 1701
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4899 CAGGATGCCGCCTTTCACTAAAATGTTTCTAGCCTATCAAGATCAGAGTTTCGACATTCC 4958

```

Qy	1702	CGACCGGCAGTTCCACTCGGTGGCGGCAGCCTGGCAG---GCACGCCTGGAGAGCCCTGC	1758
Db	4959	AGACCGAACATTTTCATTCTACAAACACAACCTGGCGCCTCTCCTCCTTTGAGTCCATGAC	5018
Qy	1759	CGATGTGAAGGAGCTCATCCCGGAATTCTTCTACTTTCTGACTTCTCTGGAGAACCAGAA	1818
Db	5019	TGATGTGAAGGAGCTGATTCCAGAGTTTTTCTATCTTCTGAGTTCTTAGTGAACCGTGA	5078
Qy	1819	CGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGTAGGCGATGTGGTGCTACC	1878
Db	5079	AGGCTTTGACTTCGGTGTTTCGTGAGAATGGAGAGCGGGTTAACCACGTCAATCTTCTCTCC	5138
Qy	1879	CCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACCGCCAGGCTCTGGAGTCGGA	1938
Db	5139	CTGGGCACGCAACGATCCTCGGCTGTTTCATCCTTATTACCGGCAAGCACTAGAGTCTGA	5198
Qy	1939	GTATGTGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGCTACAAGCAGCGGGG	1998
Db	5199	CCATGTGTCCCAGAACATCTGTCACTGGATCGACTTAGTGTTTGGCTACAAGCAAAGGG	5258
Qy	1999	GCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGAGGGGGCTGTAGA	2058
Db	5259	GAAGGCGTCTGTTCAAGCCATCAATGTCTTCCACCCTGCTACATATTTTGG---AATGGA	5315
Qy	2059	CCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTATCAGCAACTT	2118
Db	5316	TGTCTCTGCAGTTGAAGATCCAGTGCAGAGACGGGCTTTAGAAACCATGATAAAAACCTA	5375
Qy	2119	TGGGCAGACTCCCTGTCTAGCTGCTGAAGGAGCCACATCCAACCTCGGCTCTCAGCTGAG	2176
Db	5376	CGGGCAGACCCACGTCAGTTGTTCCACACAGCCCATGCCAGCCGACCTGGAGCCAAG	5433

SEQUENCE COMPARISON US 6,607,879

Part of 09/990,437 page 13

; Sequence 827, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; INFORMATION FOR SEQ ID NO: 827:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1595 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGAST01
; CLONE: 873034
US-09-023-655-827

Query Match 38.7%; Score 1367.2; DB 4; Length 1595;
Best Local Similarity 93.6%; Pred. No. 4.2e-304;
Matches 1494; Conservative 0; Mismatches 10; Indels 92; Gaps 3;

Qy 1807 GGAGAACCAGAACGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGTAGGCCGA 1866
| | | | |
Db 1 GGAGAACCAGAACGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGTAGGCCGA 60

Qy 1867 TGTGGTGCTACCCCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACC GCCAGGC 1926
| | | | |
Db 61 TGTGGTGCTACCCCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACC GCCAGGC 120

Qy 1927 TCTGGAGTCGGAGTATGTGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGCTA 1986
| | | | |
Db 121 TCTGGAGTCGGAGTATGTGTCTGCACANCTACACGAGTGGATCGACCTCATCTTTGGNTA 180

Qy 1987 CAAGCAGCGGGGGCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA 2046
| | | | |
Db 181 TAAGCAGCGGGGGCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA 240

Qy 2047 GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCAT 2106
| | | | |
Db 241 GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCAT 300

Qy 2107 TATCAGCAACTTTGGGCAGACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACTCGGCT 2166
| | | | |
Db 301 TATCAGCAACTTTGGGCAGACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACTCGGCT 360

Qy 2167 CTCAGCTGAGGAAGCAGCCCATCGCCTTGACGCCTGGACACTAACTCACCTAGCATCTT 2226
| | | | |
Db 361 CTCAGCTGAGGAAGCAGCCCATCGCCTTGACGCCTGGACACTAACTCACCTAGCATCTT 420

Qy 2227 CCAGCACCTGGACGAACTCAAGGCATTCTTCGCAGA----- 2262
| | | | |
Db 421 CCAGCACCTGGACGAACTCAAGGCATTCTTCGCAGAGGTTGTCAGTGATGGTGTACCCCT 480

Qy 2263 ----- 2262

Db 481 GGTGCTAGCCCTGGTCCCCCACCAGGAGCCCCACTCCTTCATCACCCAGGGTTCCCCAGA 540

Qy 2263 -----GGTGAAGTGTGAGTGCCAGTGGGCTGCTGGGCACCCACAGCT-GGTTGCCCTATG 2315
|||||

Db 541 CCTGTTGGTGAAGTGTGAGTGCCAGTGGGCTGCTGGGCACCCACAGCTGGGTTGCCCTATG 600

Qy 2316 ACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCACCATGGGCAGCCACAAGA 2375
|||||

Db 601 ACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCACCATGGGCAGCCACAAGA 660

Qy 2376 CGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGTGAGTGGACAAGCACTGG 2435
|||||

Db 661 CGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGTGAGTGGACAAGCACTGG 720

Qy 2436 CAGTGGCCCCGGATGGAAAGCTGCTATTTCAGCGGTGGCCACTGGGATGGCAGCCTGCGGG 2495
|||||

Db 721 CAGTGGCCCCGGATGGAAAGCTGCTATTTCAGCGGTGGCCACTGGGATGGCAGCCTGCGGG 780

Qy 2496 TGAAGTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCACCTTGATGTAGTAA 2555
|||||

Db 781 TGAAGTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCACCTTGATGTAGTAA 840

Qy 2556 CCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCGGGACACCACGT 2615
|||||

Db 841 CCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCGGGACACCACGT 900

Qy 2616 GCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCACCAAAGCCTG 2675
|||||

Db 901 GCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCACCAAAGCCTG 960

Qy 2676 TGCAGGTCTGTATGGGCATGGGGCTGCAGTGAAGCTGTGTGGCCATCAGCACTGAAGTTG 2735
|||||

Db 961 TGCAGGTCTGTATGGGCATGGGGCTGCAGTGAAGCTGTGTGGCCATCAGCACTGAAGTTG 1020

Qy 2736 ACATGGCTGTGTCTGGATCTGAGGATGGAAGTGTGATCATAACACTGTACGCCGCGGAC 2795
|||||

Db 1021 ACATGGCTGTGTCTGGATCTGAGGATGGAAGTGTGATCATAACACTGTACGCCGCGGAC 1080

Qy 2796 AGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCTGGACCTATTTTCCACCTGG 2855
|||||

Db 1081 AGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCTGGACCTATTTTCCACCTGG 1140

Qy 2856 CATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTTGGGGCCC 2915
|||||

Db 1141 CATTGGGGTCCGAATGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTTGGGGCCC 1200

Qy 2916 AGGTCACCTACTCCTTGACCTGTATTTCAGTCAATGGGAAGTTGCGGGCTTCACTGCCCC 2975
|||||

Db 1201 AGGTCACCTACTCCTTGACCTGTATTTCAGTCAATGGGAAGTTGCGGGCTTCACTGCCCC 1260

Qy 2976 TGGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCC 3035
|||||

Db 1261 TGGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCC 1320

Art Unit: 1646

SEQUENCE COMPARISON GenBank AB011112

Part of 09/990,437 page 16

AC 060288;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE KIAA0540 protein (Fragment).
 GN KIAA0540.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N.,
 RA Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 DR EMBL; AB011112; BAA25466.3; - .
 DR InterPro; IPR000409; Beige_BEACH.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF02138; Beach; 1.
 DR Pfam; PF00400; WD40; 2.
 DR ProDom; PD007848; Beige_BEACH; 1.
 DR SMART; SM00320; WD40; 3.
 DR PROSITE; PS50197; BEACH; 1.
 DR PROSITE; PS50082; WD_REPEATS_2; 2.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT NON_TER 1 1
 SQ SEQUENCE 2041 AA; 224474 MW; CC523F1D5D041F51 CRC64;

Query Match 99.5%; Score 5222; DB 4; Length 2041;
 Best Local Similarity 97.1%; Pred. No. 0;
 Matches 1003; Conservative 0; Mismatches 0; Indels 30; Gaps 1;

Qy 1 MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLE 60
 |||
 Db 1009 MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLE 1068
 Qy 61 GLRYTAVLKQQTQHSMAALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLK 120
 |||
 Db 1069 GLRYTAVLKQQTQHSMAALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLK 1128
 Qy 121 LVPNHFFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDEL 180
 |||
 Db 1129 LVPNHFFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDEL 1188
 Qy 181 AELETPM EAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEG 240
 |||
 Db 1189 AELETPM EAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEG 1248

Qy	241	IGYDFRRPLAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQ	300
Db	1249	IGYDFRRPLAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQ	1308
Qy	301	PGPIPPHTQVRNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQ	360
Db	1309	PGPIPPHTQVRNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQ	1368
Qy	361	LNTIAGRTYNDLSQYPVFPWVLQDYVSPITDLNSPAVFRDLSKPIGVVNPKHAQLVREKY	420
Db	1369	LNTIAGRTYNDLSQYPVFPWVLQDYVSPITDLNSPAVFRDLSKPIGVVNPKHAQLVREKY	1428
Qy	421	ESFEDPAGTIDKFHYGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAA	480
Db	1429	ESFEDPAGTIDKFHYGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAA	1488
Qy	481	WQARLESPADVKELEIPEFFYFPDFLENQNGFDLGLCLTNEKVGDDVLPWPWASSPEDFIQ	540
Db	1489	WQARLESPADVKELEIPEFFYFPDFLENQNGFDLGLCLTNEKVGDDVLPWPWASSPEDFIQ	1548
Qy	541	QHRQALESEYVSAHLHEWIDLIFGYKQQRGPAEEALNVFYICYEGAVDLDHVTDERERK	600
Db	1549	QHRQALESEYVSAHLHEWIDLIFGYKQQRGPAEEALNVFYICYEGAVDLDHVTDERERK	1608
Qy	601	ALEGIISNFGQTPCQLLKEPHPTRLSAEEAAHRLARLDTNPSIFQHLDLKAFFAE---	657
Db	1609	ALEGIISNFGQTPCQLLKEPHPTRLSAEEAAHRLARLDTNPSIFQHLDLKAFFAEVVS	1668
Qy	658	-----VTVSASGLLGTHSWLPYDRNISNYFSFSKDDPTM	690
Db	1669	DGVPLVLALVPHRQPHSFITQGSDDLVTVSASGLLGTHSWLPYDRNISNYFSFSKDDPTM	1728
Qy	691	GSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGDSLRTALPRGKLLSQLSCH	750
Db	1729	GSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGDSLRTALPRGKLLSQLSCH	1788
Qy	751	LDVVTCLALDTCGIYLISSSRDT TCMVWRLHQQGLSVGLAPKPVQVLYGHGAAVSCVAI	810
Db	1789	LDVVTCLALDTCGIYLISSSRDTTCMVWRLHQQGLSVGLAPKPVQVLYGHGAAVSCVAI	1848
Qy	811	STELDMAVSGSEDGTVI IHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSAW	870
Db	1849	STELDMAVSGSEDGTVI IHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSAW	1908
Qy	871	RPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA	930
Db	1909	RPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA	1968
Qy	931	PPLPMKVAIRSVAVTKERSHVLVGLLEDGKLI VVVAGQPSEVRSSQFARKLWRRSSRRISQV	990
Db	1969	PPLPMKVAIRSVAVTKERSHVLVGLLEDGKLI VVVAGQPSEVRSSQFARKLWRRSSRRISQV	2028
Qy	991	SSGETEYNPTPEAR	1003
Db	2029	SSGETEYNPTPEAR	2041

LOCUS AB011112 6364 bp mRNA linear PRI 05-JUL-2002
 DEFINITION Homo sapiens mRNA for KIAA0540 protein, partial cds.
 ACCESSION AB011112
 VERSION AB011112.3 GI:20521076
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
 Nomura,N. and Ohara,O.
 TITLE Prediction of the coding sequences of unidentified human genes. IX.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro
 JOURNAL DNA Res. 5 (1), 31-39 (1998)
 MEDLINE 98290545
 PUBMED 9628581
 REFERENCE 2 (bases 1 to 6364)
 AUTHORS Ohara,O., Nagase,T. and Ishikawa,K.
 TITLE Direct Submission
 JOURNAL Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)
 COMMENT On May 9, 2002 this sequence version replaced gi:6635198.
 FEATURES
 Location/Qualifiers
 source 1. .6364
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="pg00116"
 /sex="male"
 /tissue_type="brain"
 /clone_lib="pBluescriptII SK plus"
 /note="This sequence was replaced that of hg04185 cDNA as
 a representative cDNA sequence for KIAA0540."
 gene 1. .6364
 /gene="KIAA0540"
 CDS <1. .6128
 /gene="KIAA0540"
 /note="Start codon is not identified"
 /codon_start=3
 /product="KIAA0540 protein"
 /protein_id="BAA25466.3"
 /db_xref="GI:20521077"
 /translation="PMELRHLLRPRPGLDSEPGGAEAGKARHAGAVIRTLSGMARHQG
 PARALRYFDLTSPMAGIMVPPVQRWPGPGFTFHAWLCLHPMDTAPTAPTRPLQRKQL
 YSFFTSSGSGFEAFFTAAGTLVVAVCTRKEYLTMSLPEVSFADSAWHCVAIVHVPGRR
 PFSQNLVHVYKDGHLVKTAPLRCPSLSEPFSSCCIGSAGYRTTTTTTGLPTPPVPATL
 AYTHPALTRSQSVPASTGLGWGSGLVAPLQEGSIDSTLAGTQDTRWGSPTSLEGELGA
 VAIFHEALQATALRTLCTLGPNETAPFKPEGELHELSTRLLLHYSPOACKNNICLDLS
 PSHGLDGRLTGHRVETWDVKDVNVCVGGMGALLPLLERVAAQPKEAEAGPAETHDLVG
 PELTSGHNTQGLVPLGKSSEERMERNAVAFLMLRNFLQGHEMVNQESLVQCQGP
 IIGALLRKVPSWAMDMNVLMSAQLLMEQVAAEGSGPLLYLLYQHLLFNHFLWTLSDFAV
 RLGHIQYMSSIVREHRQKLKKYGVQFILDALRTHYSPQERERPLAADLRTVQTSLLG
 LAREFLVRSLSADDVQVTQTMLSFLAATGDDGQILRRLQQNERLPERSRQRLRLRECG

LQGLVACLPEGTVSPQLCQGLYKFLGAGSGNTAGGGGSSGTLTPASQPGTPSPLDGP
 RPFPAAPGRHSSSLSNVLEDGSLPEPTISGDDTSNTSNPQQTSEEEELCNLLTNVLFVS
 TWRGVEGSDEAAWRERQVFSVLTQLGASATLVRPPDCIKRSLEMMLESALTDIKEA
 PVGVLASLTQQALWLLRLQLDFLCAEGHGNQELWSEKLFEGVCSLLDRLGAWPHLANG
 TADLREMAQIGRLVLGYILLEDPQLHAQAYVRLHMLLQTAVPARREEACYVLSKLEA
 ALGRVLNTSSLESATDEAGSPAAAAAAAAAERCSWLVLVRLTLLDRAYEPLGLQWGL
 PSLPPTNGSPFTFFEDFQAFCATPEWRHFIDKQVQPTMSQFEMDTYAKSHDLMSGFWNA
 CYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGLRYTAVLKQQATQHSMA
 HWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVNHHFDPHLEASALRD
 NLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAELTPMEAAELDEQR
 EKLVLSAECQLVTVVAVVPGLEVTQNVFYFDGSTERVETEEGIGYDFRRPLAQLRE
 VHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQVRN
 QVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYN
 DLSQYFVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVNPKHAQLVREKYESFEDPAG
 TIDKFHYGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDSCDRQFHSVAAAWQARLE
 SPADVKELIPEFFFYFPDFLENQNGFDLGCLQLTNEKVGDDVLPWPWASSPEDFIQQHRQ
 ALESEYVSAHLHEWIDLIFGYKQRGPAEEALNVFYCYTYEYAVDLHDVTDERERKAL
 EGIISNFGQTPCQLLKEPHPTRLSAEEAAHRLARLDTNPSIFQHLDELKAFFAEVVS
 DGVPLVLALVPHRQPHSFITQGSPDLLVTVSASGLLGTHSWLPYDRNISNYFSFSKDP
 TMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGDSLRTALPRGKLLSQ
 LSCHLDVVTCLALDTCGIYLI SGRDTCMVWRLLHQGGLSVGLAPKPVQVLYGHGAA
 VSCVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPFIHFLALGSEGQI
 VVQSSAWERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHI
 LQLNTLLPAAPPLPMKVAIRSVAVTKERSHVLVGLLEDGKLIVVAGQPSEVRSSQFAR
 KLWRSSRRISQVSSGETEYNPTAR"

ORIGIN

Query Match 96.0%; Score 3391.2; DB 9; Length 6364;
 Best Local Similarity 97.4%; Pred. No. 0;
 Matches 3504; Conservative 0; Mismatches 3; Indels 91; Gaps 2;

Qy	25	GCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGGCAGGGTCCCCACTTGCAGCTGC	84
Db	2759	GCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGGCAGGGTCCCCACTTGCAGCTGC	2818
Qy	85	AGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTGCCACTGGTGCGCACGCTGCT	144
Db	2819	AGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTACCACTGGTGCGCACGCTGCT	2878
Qy	145	AGACCGTGCCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCCCTCCCTGCCACCCACCAA	204
Db	2879	AGACCGTGCCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCCCTCCCTGCCACCCACCAA	2938
Qy	205	TGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCACACCCGAATGGCGCCA	264
Db	2939	TGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCACACCCGAATGGCGCCA	2998
Qy	265	CTTCATCGACAAACAGGTACAGCCAACCATGTCCCAGTTCGAAATGGACACGTATGCTAA	324
Db	2999	CTTCATCGACAAACAGGTACAGCCAACCATGTCCCAGTTCGAAATGGACACGTATGCTAA	3058
Qy	325	GAGCCACGACCTTATGTGAGGTTTCTGGAATGCCTGCTATGACATGCTTATGAGCAGTGG	384
Db	3059	GAGCCACGACCTTATGTGAGGTTTCTGGAATGCCTGCTATGACATGCTTATGAGCAGTGG	3118
Qy	385	GCAGCGGCGCCAGTGGGAGCGCGCCAGAGTCGTCGGGCCTTCCAGGAGCTGGTGCTGGA	444

Db 3119 GCAGCGGCGCCAGTGGGAGCGCGCCAGAGTCGTCGGGCCTTCCAGGAGCTGGTGCTGGA 3178
 Qy 445 ACCTGCGCAGAGGCGGGCGCGCCTGGAGGGGGCTACGCTACACGGCAGTGCTGAAGCAGCA 504
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3179 ACCTGCGCAGAGGCGGGCGCGCCTGGAGGGGGCTACGCTACACGGCAGTGCTGAAGCAGCA 3238
 Qy 505 GGCAACGCAGCACTCCATGGCCCTGCTGCACTGGGGGGCGCTGTGGCGCCAGCTCGCCAG 564
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3239 GGCAACGCAGCACTCCATGGCCCTGCTGCACTGGGGGGCGCTGTGGCGCCAGCTCGCCAG 3298
 Qy 565 CCCATGTGGGGCCTGGGCGCTGAGGGGACACTCCCATCCCCCGCTGGAAACTGTCCAGCGC 624
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3299 CCCATGTGGGGCCTGGGCGCTGAGGGGACACTCCCATCCCCCGCTGGAAACTGTCCAGCGC 3358
 Qy 625 CGAGACATATTCACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCTCACCT 684
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3359 CGAGACATATTCACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCTCACCT 3418
 Qy 685 GGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCTGACACCCACCGAGGAGGC 744
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3419 GGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCTGACACCCACCGAGGAGGC 3478
 Qy 745 CTCCTGCTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCACCCACCCGAGTTGCTGCA 804
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3479 CTCCTGCTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCACCCACCCGAGTTGCTGCA 3538
 Qy 805 GGAGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGA 864
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3539 GGAGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGA 3598
 Qy 865 ACTGGATGAGCAGCGTGAGAAGCTGGTGCTGTGCGCCGAGTGCCAGCTGGTGACGGTAGT 924
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3599 ACTGGATGAGCAGCGTGAGAAGCTGGTGCTGTGCGCCGAGTGCCAGCTGGTGACGGTAGT 3658
 Qy 925 GGCCGTGGTCCCAGGGCTGCTGGAGGTCAACACACAGAATGTATACTTCTACGATGGCAG 984
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3659 GGCCGTGGTCCCAGGGCTGCTGGAGGTCAACACACAGAATGTATACTTCTACGATGGCAG 3718
 Qy 985 CACTGAGCGCGTGGAACCGAGGAGGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCCA 1044
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3719 CACTGAGCGCGTGGAACCGAGGAGGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCCA 3778
 Qy 1045 GCTGCGTGAGGTCCACCTGCGGCGTTTCAACCTGCGCCGTTTCAAGCTTCTTCTT 1104
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3779 GCTGCGTGAGGTCCACCTGCGGCGTTTCAACCTGCGCCGTTTCAAGCTTCTTCTT 3838
 Qy 1105 TATCGATCAGGCCAACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCAGTCTC 1164
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3839 TATCGATCAGGCCAACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCAGTCTC 3898
 Qy 1165 ATCTCCTAGCCAGACTCCGAGACCCAGCCTGGCCCCATCCACCCCATACCCAGGTACG 1224
 |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
 Db 3899 ATCTCCTAGCCAGACTCCGAGACCCAGCCTGGCCCCATCCACCCCATACCCAGGTACG 3958
 Qy 1225 GAACCAGGTGTACTCGTGGCTCCTGCGCCTACGGCCCCCCTCTCAAGGCTACCTAAGCAG 1284
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3959 GAACCAGGTGTACTCGTGGCTCCTGCGCCTACGGCCCCCCTCTCAAGGCTACCTAAGCAG 4018

Qy	1285	CCGCTCCCCCAGGAGATGCTGCGTGCCTCAGGCCCTTACCCAGAAATGGGTACAGCGTGA	1344
Db	4019	 CCGCTCCCCCAGGAGATGCTGCGTGCCTCAGGCCCTTACCCAGAAATGGGTACAGCGTGA	4078
Qy	1345	GATATCCAACCTTCGAGTACTTGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGA	1404
Db	4079	 GATATCCAACCTTCGAGTACTTGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGA	4138
Qy	1405	CCTGTCTCAGTACCCTGTGTTCCCTGGGTCCTGCAGGACTACGTGTCCCCAACCCCTGGA	1464
Db	4139	 CCTGTCTCAGTACCCTGTGTTCCCTGGGTCCTGCAGGACTACGTGTCCCCAACCCCTGGA	4198
Qy	1465	CCTCAGCAACCCAGCCGTCTTCCGGGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAA	1524
Db	4199	 CCTCAGCAACCCAGCCGTCTTCCGGGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAA	4258
Qy	1525	GCATGCCCAGCTCGTGAGGGAGAAGTATGAAAGCTTTGAGGACCCAGCAGGGACCATTGA	1584
Db	4259	 GCATGCCCAGCTCGTGAGGGAGAAGTATGAAAGCTTTGAGGACCCAGCAGGGACCATTGA	4318
Qy	1585	CAAGTTCCACTATGGCACCCACTACTCCAATGCAGCAGGCGTGATGCACTACCTCATCCG	1644
Db	4319	 CAAGTTCCACTATGGCACCCACTACTCCAATGCAGCAGGCGTGATGCACTACCTCATCCG	4378
Qy	1645	CGTGGAGCCCTTCACCTCCCTGCACGTCCAGCTGCAAAGTGGCCGCTTTGACTGCTCCGA	1704
Db	4379	 CGTGGAGCCCTTCACCTCCCTGCACGTCCAGCTGCAAAGTGGCCGCTTTGACTGCTCCGA	4438
Qy	1705	CCGGCAGTTCCAACCTCGGTGGCGGCAGCCTGGCAGGCACGCCTGGAGAGCCCTGCCGATGT	1764
Db	4439	 CCGGCAGTTCCAACCTCGGTGGCGGCAGCCTGGCAGGCACGCCTGGAGAGCCCTGCCGATGT	4498
Qy	1765	GAAGGAGCTCATCCCGGAATTCTTCTACTTTCTGACTTCCTGGAGAACCAGAACGGTTT	1824
Db	4499	 GAAGGAGCTCATCCCGGAATTCTTCTACTTTCTGACTTCCTGGAGAACCAGAACGGTTT	4558
Qy	1825	TGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGTAGGCGATGTGGTGCTACCCCCGTG	1884
Db	4559	 TGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGTAGGCGATGTGGTGCTACCCCCGTG	4618
Qy	1885	GGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACCGCCAGGCTCTGGAGTCGGAGTATGT	1944
Db	4619	 GGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACCGCCAGGCTCTGGAGTCGGAGTATGT	4678
Qy	1945	GTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGCTACAAGCAGCGGGGGCCAGC	2004
Db	4679	 GTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGCTACAAGCAGCGGGGGCCAGC	4738
Qy	2005	CGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGAGGGGGCTGTAGACCTGGA	2064
Db	4739	 CGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGAGGGGGCTGTAGACCTGGA	4798
Qy	2065	CCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTATCAGCAACTTTGGGCA	2124
Db	4799	 CCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTATCAGCAACTTTGGGCA	4858

Qy	2125	GACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACCTCGGCTCTCAGCTGAGGAAGCAGC	2184
Db	4859	GACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACCTCGGCTCTCAGCTGAGGAAGCAGC	4918
Qy	2185	CCATCGCCTTGACAGCCTGGACACTAACTCACCTAGCATCTTCCAGCACCTGGACGAACT	2244
Db	4919	CCATCGCCTTGACAGCCTGGACACTAACTCACCTAGCATCTTCCAGCACCTGGACGAACT	4978
Qy	2245	CAAGGCATTCTTCGCAGA-----	2262
Db	4979	CAAGGCATTCTTCGCAGAGGTTGTCAGTGATGGTGTACCCCTGGTGCTAGCCCTGGTCCC	5038
Qy	2263	-----GGTGACTGTGAG	2274
Db	5039	CCACCGGCAGCCCCACTCCTTCATCACCCAGGGTCCCCAGACCTGTTGGTGACTGTGAG	5098
Qy	2275	TGCCAGTGGGCTGCTGGGCACCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTA	2334
Db	5099	TGCCAGTGGGCTGCTGGGCACCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTA	5158
Qy	2335	CTTCAGCTTCAGCAAAGACCCACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGG	2394
Db	5159	CTTCAGCTTCAGCAAAGACCCACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGG	5218
Qy	2395	CCCGTGGGTGCCAGGCAGTGGTGTGAGTGGACAAGCACTGGCAGTGGCCCCGGATGGAAA	2454
Db	5219	CCCGTGGGTGCCAGGCAGTGGTGTGAGTGGACAAGCACTGGCAGTGGCCCCGGATGGAAA	5278
Qy	2455	GCTGCTATTTCAGCGGTGGCCACTGGGATGGCAGCCTGCGGGTGACTGCACTACCCCGTGG	2514
Db	5279	GCTGCTATTTCAGCGGTGGCCACTGGGATGGCAGCCTGCGGGTGACTGCACTACCCCGTGG	5338
Qy	2515	CAAGCTGTTGAGCCAGCTCAGCTGCCACCTTGATGTAGTAACCTGCCTTGCACTGGACAC	2574
Db	5339	CAAGCTGTTGAGCCAGCTCAGCTGCCACCTTGATGTAGTAACCTGCCTTGCACTGGACAC	5398
Qy	2575	CTGTGGCATCTACCTCATCTCAGGCTCCCGGGACACCACGTGCATGGTGTGGCGGCTCCT	2634
Db	5399	CTGTGGCATCTACCTCATCTCAGGCTCCCGGGACACCACGTGCATGGTGTGGCGGCTCCT	5458
Qy	2635	GCATCAGGGTGGTCTGTGCTAGTAGGCCTGGCACCAAAGCCTGTGCAGGTCCTGTATGGGCA	2694
Db	5459	GCATCAGGGTGGTCTGTGCTAGTAGGCCTGGCACCAAAGCCTGTGCAGGTCCTGTATGGGCA	5518
Qy	2695	TGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACTGAACTTGACATGGCTGTGTCTGGATC	2754
Db	5519	TGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACTGAACTTGACATGGCTGTGTCTGGATC	5578
Qy	2755	TGAGGATGGAACCTGTGATCATAACACTGTACGCCGCGGACAGTTTGTAGCGGCACTACG	2814
Db	5579	TGAGGATGGAACCTGTGATCATAACACTGTACGCCGCGGACAGTTTGTAGCGGCACTACG	5638
Qy	2815	GCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACCTGGCATTGGGGTCCGAAGGCCA	2874
Db	5639	GCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACCTGGCATTGGGGTCCGAAGGCCA	5698
Qy	2875	GATTGTGGTACAGAGCTCAGCGTGGGAACGTCTCTGGGGCCCAGGTCACCTACTCCTTGCA	2934

Db 5699 GATTGTGGTACAGAGCTCAGCGTGGGAACGTCTCTGGGGCCCAGGTCACCTACTCCTTGCA 5758
Qy 2935 CCTGTATTCACTCAATGGGAAGTTGCGGGCTTCACTGCCCCCTGGCAGAGCAGCCTACAGC 2994
Db 5759 CCTGTATTCACTCAATGGGAAGTTGCGGGCTTCACTGCCCCCTGGCAGAGCAGCCTACAGC 5818
Qy 2995 CCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCCAGTGCGCCCTGCACATCCT 3054
Db 5819 CCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCCAGTGCGCCCTGCACATCCT 5878
Qy 3055 CCAACTAAACACACTGCTCCCGGCCGCGCCTCCCTTGCCCATGAAGGTGGCCATCCGCAG 3114
Db 5879 CCAACTAAACACACTGCTCCCGGCCGCGCCTCCCTTGCCCATGAAGGTGGCCATCCGCAG 5938
Qy 3115 CGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCCTGGAGGATGGCAAGCTCAT 3174
Db 5939 CGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCCTGGAGGATGGCAAGCTCAT 5998
Qy 3175 CGTGGTGGTTCGCGGGGCAGCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGCGGAAGCTGTG 3234
Db 5999 CGTGGTGGTTCGCGGGGCAGCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGCGGAAGCTGTG 6058
Qy 3235 GCGGTCTCTCGCGGCGCATCTCCCAGGTGTCTCTCGGGAGAGACGGAATACAACCCTACTGA 3294
Db 6059 GCGGTCTCTCGCGGCGCATCTCCCAGGTGTCTCTCGGGAGAGACGGAATACAACCCTACTGA 6118
Qy 3295 GGC GCGCTGAACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCCCGGCAGGCCTGGCCCCGG 3354
Db 6119 GGC GCGCTGAACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCCCGGCAGGCCTGGCCCCGG 6178
Qy 3355 GAGGCCCCGCCCAGAAGTCGGCGGGAACACCCCGGGGTGGGCAGCCCA-GGGGGTGAGCG 3413
Db 6179 GAGGCCCCGCCCAGAAGTCGGCGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGAGCG 6238
Qy 3414 GGGCCCACCCTGCCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCGG 3473
Db 6239 GGGCCCACCCTGCCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCGG 6298
Qy 3474 GCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGCGCCCTGAGGGCCAGCACTGGCGTCT 3531
Db 6299 GCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGCGCCCTGAGGGCCAGCACTGGCGTCT 6356